Seed image analysis and taxonomy of *Diplotaxis* DC. (Brassicaceae, Brassiceae)

OSCAR GRILLO , DAVID DRAPER , GIANFRANCO VENORA & JUAN BAUTISTA MARTÍNEZ-LABORDE

The genus *Diplotaxis*, comprising 32 or 34 species, plus several additional infraspecific taxa, displays a considerable degree of heterogeneity in the morphology, molecular markers, chromosome numbers and geographical amplitude of the species. The taxonomic relationships within the genus *Diplotaxis* were investigated by phenetic characterisation of germplasm belonging to 27 taxa of the genus, because there is an increasing interest in *Diplotaxis*, since some of its species (*D. tenuifolia*, *D. muralis*) are gathered or cultivated for human consumption, whereas others are frequent arable weeds (*D. erucoides*) in many European vineyards. Using a computer-aided vision system, 33 morpho-colorimetric features of seeds were electronically measured. The data were used to implement a statistical classifier, which is able to discriminate the taxa within the genus *Diplotaxis*, in order to compare the resulting species grouping with the current infrageneric systematics of this genus. Despite the high heterogeneity of the samples, due to the great intra-population variability, the stepwise Linear Discriminant Analysis method, applied to distinguish the groups, was able to reach over 80% correct identification. The results obtained allowed us to confirm the current taxonomic position of most taxa and suggested the taxonomic position of others for reconsideration.

Key words: computer vision, germplasm characterisation, Linear Discriminant Analysis, morpho-colorimetric measurements, seed identification, statistical classification

Introduction

The genus, *Diplotaxis* DC. (Brassicaceae, tribe Brassiceae), currently comprises 32 species (Warwick *et al.*, 2006) or 34 (Table 1), plus several additional infraspecific taxa, native to Europe, the Mediterranean basin, SW Asia (up to the Himalayas) and Macaronesia. There is an increasing interest in *Diplotaxis*, since some of its species (*D. tenuifolia*, *D. muralis*) are gathered or cultivated for human consumption as rocket salad (Pignone, 1997; Pimpini & Enzo, 1997), whereas others are frequent arable weeds (*D. erucoides*) in many European vineyards (Sans & Masalles, 1994).

The genus displays a considerable degree of heterogeneity in the morphology, molecular markers, chromosome numbers and geographical amplitude of the species (Table 1). Its highest diversity is found in NW Africa and the Iberian Peninsula, where several endemic taxa are restricted to very limited areas (*D. ibicensis*, *D. brachycarpa*) or even small islands (*D. siettiana*), or occupy more extensive regions (*D. assurgens*, *D. virgata*). Other species display much larger ranges, either across N Africa and SW Asia (*D. harra*) or central and southern Europe (*D. tenuifolia*); a few of these have colonised elsewhere (N and S America, Australia, etc.).

Chromosome numbers are known for most taxa and range from n = 7 in *D. erucoides*, through n = 8, 9, 10 and 11, to n = 13 in the *D. harra* and allied species, *D. muralis*. The only species with a higher ploidy level has n = 21, and according to solid evidence (Harberd & McArthur, 1972; Sánchez-Yélamo & Martínez-Laborde, 1991; Warwick & Anderson, 1997; Eschmann-Grupe *et al.*, 2003), is an amphidiploid probably arisen from *D. tenuifolia* (n = 11) and *D. viminea* (n = 10). Morphological variation in this genus includes remarkable differences in habit (from

Table 1. *Diplotaxis* taxa according to the subgenera (SG) and sections (S) proposed in Gómez-Campo & Martínez-Laborde (1998), diploid chromosome numbers (2n) and geographical areas of distribution.

SG[a]	S[b]	Species (and subspecies)	2n[c]	Geographical area
D		*D. tenuifolia* (L.) DC.	22	Europe, Middle East
		D. cretacea Kotov	22	NE Ukraine and adjacent Russia
		D. muralis (L.) DC.		
		subsp. *muralis*	42	Europe
		subsp. *ceratophylla* (Batt.) Mart.-Laborde	-	NE Algeria
		D. scaposa DC.		Island of Lampedusa
		D. simplex (Viv.) Spr.	22	N Africa
		D. viminea (L.) DC.	20	Europe, N Africa, Middle East
		D. harra (Forssk.) Boiss.		
		subsp. *harra*	26	N Africa, Middle East
		subsp. *crassifolia* (Raf.) Maire	26	Sicily
		subsp. *lagascana* (DC.) O. Bolòs & Vigo	26	SE Spain
		D. kohlaanensis A. G. Miller & J. Nyberg	-	Yemen
		D. villosa Boulos & Jall.	-	Jordan
		D. pitardiana Maire	-	NW Africa
		D. nepalensis Hara	-	Nepal
		D. antoniensis Rustan	-	Cape Verde
		D. glauca (J. A. Schmidt) O. E. Schulz	26	Cape Verde
		D. gorgadensis Rustan		
		subsp. *gorgadensis*	-	Cape Verde
		subsp. *brochmannii* Rustan	26	Cape Verde
		D. gracilis (Webb) O. E. Schulz	26	Cape Verde
		D. hirta (A. Chev.) Rustan & Borgen	26	Cape Verde
		D. sundingii Rustan	26	Cape Verde
		D. varia Rustan	-	Cape Verde
		D. vogelli (Webb) Cout.	-	Cape Verde
H		*D. acris* (Forssk.) Boiss.	22	N Africa, Middle East (to Iraq)
		D. griffithii (Hook.f. & W. Thomps.) Boiss.	-	Afghanistan, Pakistan
R	Rh	*D. assurgens* (Delile) Grenier	18	Morocco
		D. berthautii Braun-Blanq. & Maire	18	S Morocco
		D. brachycarpa Godron	18	NE Algeria
		D. catholica (L.) DC.	18	Iberian Peninsula, N Morocco
		D. ollivieri Maire	-	S Morocco
		D. siifolia Kunze		
		subsp. *siifolia*	20	Iberian Peninsula, NW Africa
		subsp. *bipinnatifida* (Coss.) Mart.-Laborde	-	S Morocco
		subsp. *vicentina* (Sampaio) Mart.-Laborde	20	SW Portugal
		D. tenuisiliqua Delile		
		subsp. *tenuisiliqua*	18	N Morocco, NW Algeria
		subsp. *rupestris* (J. Ball) Mart.-Laborde	-	S Morocco
		D. virgata (Cav.) DC.	18	Iberian Peninsula, Morocco
	Hc	*D. ibicensis* (Pau) Gómez-Campo	16	E Spain coast, Balearic Islands
		D. brevisiliqua (Coss.) Mart.-Laborde	16	NE Morocco, NW Algeria
		D. ilorcitana (Sennen) Aedo, Mart.-Laborde & Muñoz Garm.	16	E Spain
		D. siettiana Maire	16	Island of Alboran
	Hp	*D. erucoides* (L.) DC.		
		subsp. *erucoides*	14	Europe, N Africa, Middle East
		subsp. *longisiliqua* (Coss.) Gómez-Campo	14	NE Algeria

[a]Subgenera: D, *Diplotaxis*; H, *Hesperidium* (O. E. Schulz) Nègre; R, *Rhynchocarpum* (Prantl) Mart.-Laborde.
[b]Sections: Rh, *Rhynchocarpum*; Hc, *Heterocarpum* Mart.-Laborde; Hp, *Heteropetalum* Mart.-Laborde.
[c]As reported by Amin (1972), Harberd (1972, 1976), Gómez-Campo (1980), Takahata & Hinata (1978), Martínez-Laborde (1988, 1991), Fernandes & Queirós (1970–71) and Rustan (1996).

annuals to subshrubby perennials), petal shape (with a distinct claw or a tapering limb), colour (mostly yellow, but also white or violet) and venation (brochidodromous or cladodromous to eucamptodromous) and fruit structure. As in most other Brassiceae, a number of *Diplotaxis* species are heteroarthrocarpous (presence of seeds in the stylar portion of fruit). On the basis of this variation, the latest infrageneric system, proposed by Gómez-Campo

& Martínez-Laborde (1998), recognises three subgenera: *Diplotaxis*, *Hesperidium* and *Rhynchocarpum*, the latter including three sections, namely *Rhynchocarpum*, *Heteropetalum* and *Heterocarpum* (Table 1). The system turns out to be rather consistent with variation in known chromosome numbers and cytodemes (Harberd, 1972, 1976; Takahata & Hinata, 1983; Prakash *et al.*, 1999). The molecular evidence also indicates that *Diplotaxis* constitutes a remarkably heterogeneous, even polyphyletic genus. Phylogenetic analyses of chloroplast-DNA restriction site variation of 21 *Diplotaxis* taxa (Warwick *et al.*, 1992) showed a clear separation into the two clades previously established for the whole tribe and designated as the Rapa-Oleracea and Nigra lineages by Warwick & Black (1991) or the Brassica and Sinapis lineages by Pradhan *et al.* (1992). The Rapa-Oleracea lineage includes the taxa belonging to subgen. *Diplotaxis*, plus both subspecies of *D. erucoides* (subgen. *Rhynchocarpum* sect. *Heteropetalum*), whereas all the remaining studied taxa belonging to subgen. *Rhynchocarpum* appear in the Nigra lineage (no taxon of subgen. *Hesperidium* has been so far examined for molecular markers). Grouping within each lineage is rather consistent with known chromosome numbers and cytodemes as recognised by Harberd (1976) and Takahata & Hinata (1983). In the Rapa-Oleracea lineage, one clade includes two sister groups, one with *D. harra* (n = 13) and the other with *D. tenuifolia* and related taxa (n = 11), while *D. muralis* and *D. viminea* appear in a third clade. A fourth clade corresponds to subgen. *Rhynchocarpum* sect. *Heteropetalum*. In the Nigra lineage, one clade contains most species of sect. *Rhynchocarpum*, though *D. brachycarpa* appears in a second clade, while a third clade corresponds exactly to sect. *Heterocarpum*. The dendrogram obtained by Martín & Sánchez-Yélamo (2000) on the basis of nuclear DNA microsatellite markers of 10 *Diplotaxis* species shows two main branches, approximately corresponding to the mentioned lineages, although one of the branches combines species of both lineages. The phenogram obtained by Eschmann-Grupe *et al.* (2003) with RAPD data of 18 *Diplotaxis* species is more ladder-like in structure and therefore the separation of taxa into the two lineages is less clear-cut, but still, clusters correspond quite well to within-lineage clades in Warwick *et al.* (1992) and to known chromosome numbers and cytodemes.

The seed morphology of this genus has received little attention to date. Bengoechea & Gómez-Campo (1975) included 15 *Diplotaxis* taxa in their comprehensive survey of seed exomorphology and anatomy in the Brassiceae. Martínez-Laborde (1988) examined a few exomorphological seed traits in 29 of the *Diplotaxis* taxa (species and subspecies) listed in Table 1. According to these authors, *Diplotaxis* seeds are ochre to brown coloured, small (0.6–1.3 mm long × 0.4–1.0 mm wide), and ovoid to ellipsoid in shape (Fig. 1). The only, notable exception is *D. siifolia*, which has more or less spherical seeds, more similar to those of *Brassica*, not only in shape, but also in the extension of the thickened portion of the radial (anticlinal) cell walls of the subepidermal, palisade layer of the testa (Bengoechea & Gómez-Campo, 1975).

In the last two decades, a remarkable increase in image analysis applications has been applied in the plant biology research field. Until recently, the dimensional measurements as length and width of the seeds were made manually, generally by calipers, while fixed categories officially recognised, reported by Martin (1946), Stearn (1980) & Werker (1997), were used to describe contour shapes. With the same principles, colour evaluation was commonly executed by comparison with the Munsell® Colour Charts (Fagundez & Izco, 2004). It is evident that there are difficulties

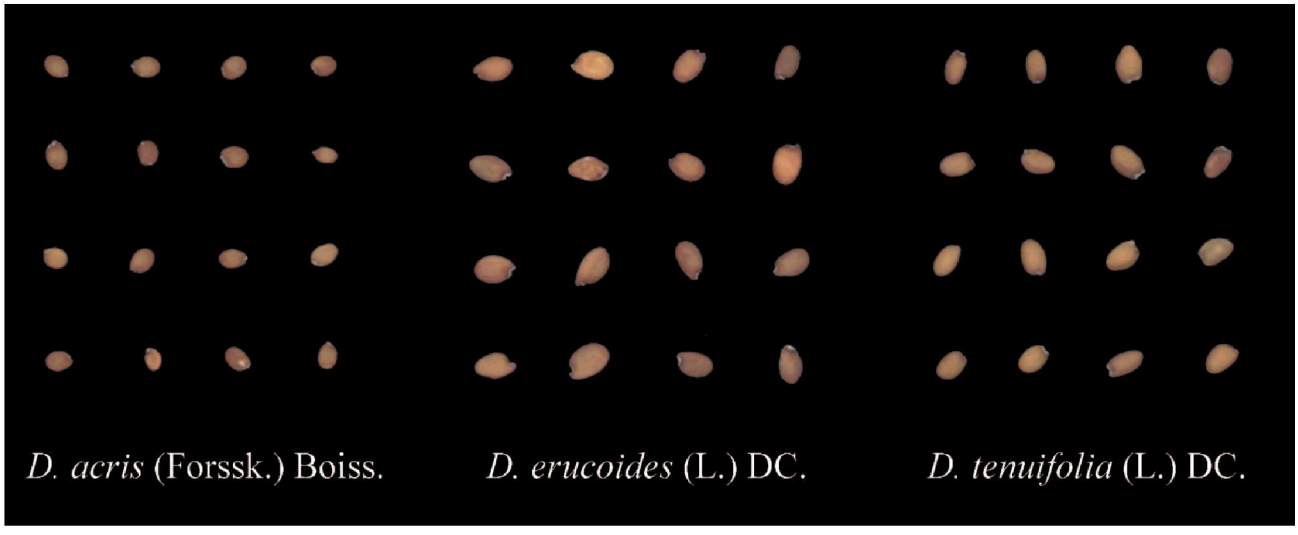

Fig. 1. Seeds of *D. acris*, *D. erucoides* and *D. tenuifolia*.

making these measurements objective and repeatable, especially when the seed dimension is extremely reduced (e.g. in *Brassicaceae*, *Cynomoriaceae*, *Primulaceae*, *Rubiaceae*, *Scrophulariaceae*, etc.). A technological evolution occurred that resulted in overcoming some of these limits (e.g. difficulty to obtain accurate, objective and repeatable results). The recent literature proves how this innovative technology improved morphometrics and colour evaluation (Liao *et al.*, 1994; Granitto *et al.*, 2003; Shahin & Symons, 2003*a*; Kiliç *et al.*, 2007; Venora *et al.*, 2007, 2009*a*, 2009*b*; Bacchetta *et al.*, 2008, 2011*a*; Dana & Ivo, 2008; Grillo *et al.*, 2010, 2011; Smykalova *et al.*, 2011). The importance of morphologic seed traits such as shape, size and external ornamentation, as diagnostic factors in plant taxonomy, is emphasized by the constantly increasing availability of seeds collected from wild plants cultivated *ex situ* (e.g. botanic gardens) and stored in germplasm banks. The implementation of a workable system, standardised on the basis of a database of morpho-colorimetric features, could be a valid tool to support the accession activities of germplasm banks, during seed cataloguing and identification or for determination and revision of critical taxonomic groups.

The aims of the present study were to use seed morpho-colorimetric data obtained by image analysis to implement a dedicated statistical classifier able to discriminate the taxa within the genus *Diplotaxis* and to compare the resulting species grouping with the current infrageneric systematics of this genus, as well as with groupings more recently revealed by molecular markers.

Materials and methods

Selected germplasm

Seed samples belonging to the Plant Genebank of the Universidad Politécnica de Madrid (BGV-UPM), and corresponding to most *Diplotaxis* taxa (27 out of the 34 species and subspecies listed in Table 1) were analysed. The accession codes of samples in the genebank and the localities of collection and the amount of the seeds investigated were recorded (Table 2). All the analysed seed accessions were stored in the same conditions and for a period longer than 15 years. This allowed us to exclude any possible variation in seed colour, due to the ageing process. In order to guarantee the representations of accessions and to minimise the intraspecific changes of shape and sizes of the seeds, due to the seed position inside the fruit and to the fruit position on the plant (Harper *et al.*, 1970), all seeds of whole accessions were measured (total of 8918 seeds analysed). Although small differences in seed size and shape can exist between different populations of the same taxon, principally due to climatic and edaphic factors as well as to

Table 2. Seed samples of *Diplotaxis* taxa (species and subspecies) investigated.

Accession code	Species and subspecies	Geographical origin of seeds	Seed amount
BGV-UPM-8860	*D. acris*	Israel (unknown locality)	222
BGV-UPM-2978	*D. assurgens*	Morocco, South of Agadir	329
BGV-UPM-7522	*D. berthautii*	Morocco, 120 km North of Marrakech	594
BGV-UPM-6467	*D. brachicarpa*	Algeria, North of Sidi Aïssa	322
BGV-UPM-7517	*D. brevisiliqua*	Morocco, Cala Iris	784
BGV-UPM-2949	*D. catholica*	Spain, Toledo, Talaverilla la Nueva	223
BGV-UPM-1445	*D. cretacea*	Unknown (Moscow Botanical Garden)	270
BGV-UPM-5056	*D. erucoides* subsp. *erucoides*	Spain, Lleida, Preixana	444
BGV-UPM-6483	*D. erucoides* subsp. *longisiliqua*	Algeria, between El Kantara and Batna	483
BGV-UPM-6472	*D. harra* subsp. *harra*	Algeria, 100 km West of Biskra	337
BGV-UPM-6635	*D. harra* subsp. *crassifolia*	Italy, Sicily, Caltanissetta	416
BGV-UPM-9047	*D. harra* subsp. *lagascana*	Spain, Alicante, La Albufereta	239
BGV-UPM-7032	*D. ibicensis*	Spain, Ibiza, Cala Eubarca	233
BGV-UPM-4065	*D. ilorcitana*	Spain, Almería, Tabernas	241
BGV-UPM-4678	*D. muralis* subsp. *muralis*	Tunisia, 11 km East of Gafsa	169
BGV-UPM-6486	*D muralis* subsp. *ceratophylla*	Aleria, Tazoult-Lambese	163
BGV-UPM-9250	*D. ollivieri*	Morocco, 10 km South of Goulimine	69
BGV-UPM-3025	*D. siettiana*	Spain, Island of Alborán	253
BGV-UPM-2964	*D. siifolia* subsp. *siifolia*	Morocco, Kenitra	250
BGV-UPM-2970	*D. siifolia* subsp. *bipinnatifida*	Morocco, Agadir	357
BGV-UPM-7620	*D. siifolia* subsp. *vicentina*	Portugal, Algarve, Aljezur	281
BGV-UPM-6453	*D. simplex*	Tunisia, North of Gafsa	397
BGV-UPM-7448	*D. tenuifolia*	Turkey, Iznik	298
BGV-UPM-7521	*D. tenuisiliqua* subsp. *tenuisiliqua*	Morocco, 75 km North of Ben Guerir	496
BGV-UPM-7527	*D. tenuisiliqua* subsp. *rupestris*	Morocco, Marrakech	473
BGV-UPM-8071	*D. viminea*	Spain, Tarragona	110
BGV-UPM-3066	*D. virgata*	Spain, Madrid	465

Table 3. List of features measured on seeds.

Feature		Description
Morphometric features		
A	Area	Seed area (mm^2)
P	Perimeter	Seed perimeter (mm)
P_{conv}	Convex perimeter	Convex perimeter of the seed (mm)
P_{Crof}	Crofton perimeter	Crofton perimeter of the seed (mm)
P_{conv}/P_{Crof}	Perimeter ratio	Ratio between convex and Crofton's perimeters
D_{max}	Max diameter	Maximum diameter of the seed (mm)
D_{min}	Min diameter	Minimum diameter of the seed (mm)
D_{min}/D_{max}	Feret ratio	Ratio between minimum and maximum diameters
Sf	Shape factor	Seed shape descriptor $= (4\pi A)/P^2$ (normalized value)
Rf	Roundness factor	Seed roundness descriptor $= (4A)/(\pi D_{max}{}^2)$ (normalized value)
Ecd	Eq. circular diameter	Diameter of a circle with equivalent area (mm)
EA_{max}	Maximum ellipse axis	Maximum axis of an ellipse with equivalent area (mm)
EA_{min}	Minimum ellipse axis	Minimum axis of an ellipse with equivalent area (mm)
Colourimetric features		
R_{mean}	Mean red channel	Red channel mean value of seed pixels (grey levels)
R_{sd}	Red std. deviation	Red channel standard deviation of seed pixels
G_{mean}	Mean green channel	Green channel mean value of seed pixels (grey levels)
G_{sd}	Green std. deviation	Green channel standard deviation of seed pixels
B_{mean}	Mean blue channel	Blue channel mean value of seed pixels (grey levels)
B_{sd}	Blue std. deviation	Blue channel standard deviation of seed pixels
H_{mean}	Mean hue channel	Hue channel mean value of seed pixels (grey levels)
H_{sd}	Hue std. deviation	Hue channel standard deviation of seed pixels
L_{mean}	Mean lightness ch.	Lightness channel mean value of seed pixels (grey levels)
L_{sd}	Lightness std. dev.	Lightness channel standard deviation of seed pixels
S_{mean}	Mean saturation ch.	Saturation channel mean value of seed pixels (grey levels)
S_{sd}	Saturation std. dev.	Saturation channel standard deviation of seed pixels
Densitometric features		
D_{mean}	Mean density	Density channel mean value of seed pixels (grey levels)
D_{sd}	Density std. deviation	Density channel standard deviation of seed pixels
S	Skewness	Asymmetry degree of intensity values distribution (grey levels)
K	Kurtosis	Peakness degree of intensity values distribution (densitometric units)
H	Energy	Measure of the increasing intensity power (densitometric units)
E	Entropy	Dispersion power (bit)
D_{sum}	Density sum	Sum of density values of the seed pixels (grey levels)
SqD_{sum}	Square density sum	Sum of the squares of density values (grey levels)

genotype–environment interactions, the differences were minimal (Bacchetta *et al.* 2011*b*); moreover, considering the specific phenotypic representation of the BGV-UPM seed accessions, the intra-population differences were not considered in this study.

Image analysis

The sample images were acquired according to Bacchetta *et al.* (2008) by a flatbed scanner (HP Scanjet 4890), with a resolution of 600 dpi and a scanning area not exceeding 2800 × 2800 pixels. Using a KS-400 V. 3.0 (Carl Zeiss, Vision, Oberkochen, Germany) image analysis system and its libraries, a specific macro was developed in the Image Analysis Laboratory of the Stazione Sperimentale di Granicoltura per la Sicilia (SSG), to obtain measurements of 33 morpho-colorimetric features of seeds (Table 3). The scanner was calibrated for colour matching following the protocol of Shahin and Symons (2003*b*) before image acquisition, as suggested by Venora *et al.* (2009*b*).

Statistical classifier

The data were evaluated statistically by applying the stepwise Linear Discriminant Analysis (LDA) algorithm, using SPSS software package release 15 (SPSS Inc. 1989–2006). This approach is commonly used to classify/identify unknown groups characterised by quantitative and qualitative variables (Fisher, 1936, 1940). The best features for seed sample identification were detected implementing a stepwise LDA method and a statistical classifier to discriminate and classify the seeds on the basis of the selected characters. When several variables are available, the stepwise method can be useful by automatically selecting the best characters on the basis of three statistical variables: Tolerance, F-to-enter and F-to-remove. The Tolerance value indicates the proportion of a variable variance not accounted for by other independent variables in the equation. A variable with very low Tolerance value provides little information to a model. F-to-enter and F-to-remove values define the power of each variable in the model and they are useful to describe what happens if a variable is inserted or removed, respectively,

from the current model. This method starts with a model that does not include any of the variables. At each step, the variable with the largest F-to-enter value that exceeds the entry criteria chosen ($F \geq 3.84$) is added to the model. The variables left out of the analysis at the last step have F-to-enter values smaller than 3.84, hence no more are added. The process was automatically stopped when no remaining variables increased the discrimination ability. Afterwards, the cross-validation procedure was applied to validate the performance of the developed classifier. This method is useful to analyse small datasets when a broad group of new unknown cases is lacking. It tests individual cases and classifies them on the basis of all others (SPSS Application Guide, 1999).

Results and discussion

The morpho-colorimetric analysis of the seed samples of *Diplotaxis* allowed us to obtain extremely precise data about their size, shape and colour (raw data not shown). Seed length and width mean values were 910.49 ± 162.92 μm and 633.67 ± 111.51 μm, respectively, with a mean diameter ratio of 0.70 ± 0.08. The shape is ovoid to ellipsoid, as determined by shape and roundness factors values (0.92 ± 0.05 and 0.64 ± 0.08, respectively). The only exception was *D. siifolia*, whose seeds are subspherical, with a diameter ratio, shape and roundness factor mean values of 0.87 ± 0.05, 0.97 ± 0.04 and 0.80 ± 0.05, respectively. These results are in accordance with previous measurements reported by Bengoechea & Gómez-Campo (1975) & Martínez-Laborde (1988).

Key parameters

Evaluating the contribution of the variables using the discrimination algorithm (LDA), it was possible to identify the features that, more than others, were relevant for the intrageneric separation of the *Diplotaxis* taxa included in this study. Considering the number of steps used by the stepwise method, Tolerance and F-to-remove values, it was possible to observe that 29 out of the 33 features used were evaluated by the statistical classifier to discriminate among the taxa (Table 4). The selected parameters were prevalently related to the colour or more in general to the densitometric features of the seeds, and six of these with highest F-to-remove values were colorimetric (Table 4). This result confirms that, apart from a few cases in which the seeds look morphologically different (*D. siettiana*, *D. siifolia* and *D. tenuifolia*), the seed shape and size are very similar in the species investigated.

Species discrimination

A fairly satisfactory level of *Diplotaxis* taxa (species and subspecies) discrimination was achieved by image analysis of morpho-colorimetric seed data. An overall cross-validation percentage of correct classification of 80.7% was reached for most taxa, with identification performance ranging from 70.8% (in *D. ibicensis*) to 97.3% (in *D. viminea*), with the only exceptions of *D. acris* (56.3%), *D. harra* subsp. *lagascana* (65.7%), *D. muralis* subsp. *cerato-phylla* (66.9%) and *D. tenuisiliqua* subsp. *rupestris* (64.5%) (Table 5).

The rather poor identification of the Saharan *D. acris* (subgen. *Hesperidium*) seeds arose mostly from mistakes for those of *D. brevisiliqua* (20.7%) and *D. siettiana* (4.1%), both in subgen. *Rhynchocarpum* and for three subspecies of *D. harra* (13.6% altogether). They are all taxa from the southern part of the geographical range of the genus with smaller seeds. Even if it was plausible to think of seed polymorphism phenomena as a cause for the reduced percentage of correct identification of *D. acris*, the relative intra-population phenotypic variability is fully considered by the LDA and included in the seed sample size. Furthermore, seed polymorphism, when it is present, would only affect a few characters that consequently are automatically rejected by the stepwise procedure as non-significatively relevant to the discrimination process. Therefore, the lower amount of available discriminant features could be the direct cause of the poor performance.

Table 4. Ranking of the selected features after 29 cycles of stepwise analysis.

Step	Feature	Tolerance	F-to-remove
1	G_{mean}	0.094	199.584
2	SqD_{sum}	0.017	158.198
3	D_{sd}	0.142	148.171
4	S_{sd}	0.221	141.389
5	G_{sd}	0.197	131.760
6	R_{sd}	0.112	105.189
7	A	0.014	105.024
8	D_{sum}	0.014	93.790
9	EA_{min}	0.002	88.927
10	B_{sd}	0.389	87.233
11	Ecd	0.001	76.720
12	EA_{max}	0.002	72.771
13	L_{sd}	0.097	72.443
14	S_{mean}	0.003	72.132
15	D_{max}	0.008	65.834
16	R_{mean}	0.034	59.820
17	Rf	0.032	59.123
18	B_{mean}	0.104	49.907
19	D_{mean}	0.003	45.596
20	E	0.246	36.601
21	H_{mean}	0.018	29.588
22	S	0.082	25.748
23	H_{sd}	0.213	24.921
24	K	0.084	23.867
25	P_{conv}	0.001	21.471
26	H	0.719	20.365
27	P_{Crof}	0.003	12.535
28	P_{conv}/P_{Crof}	0.081	6.298
29	D_{min}/D_{max}	0.100	4.156

Table 5. Cross-validated percentages of correct identification for the *Diplotaxis* taxa classifier. The number of seeds is in parentheses.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	Total
D. acris (1)	**56,3%** (**125**)				20,7% (46)					3,6% (8)	3,2% (7)	6,8% (15)	0,5% (1)				1,8% (4)	4,1% (9)	0,9% (2)	0,5% (1)		0,9% (2)	0,9% (2)					100,0% (222)
D. assurgens (2)		**82,4%** (**271**)	3,6% (12)			4,0% (13)	0,3% (1)									0,3% (1)								7,6% (25)	1,8% (6)			100,0% (329)
D. berthautii (3)		3,5% (21)	**86,5%** (**512**)			0,5% (3)										0,7% (4)								5,2% (31)	2,0% (12)	0,2% (1)	1,3% (8)	100,0% (594)
D. brachycarpa (4)				**73,3%** (**236**)	0,3% (1)			0,3% (1)	22,0% (71)		0,6% (2)				1,2% (4)	0,3% (1)							1,6% (5)			0,3% (1)		100,0% (322)
D. brevisiliqua (5)	4,2% (33)				**83,2%** (**652**)					0,1% (1)	0,5% (4)	0,9% (7)	6,0% (47)	0,5% (4)	0,4% (3)		0,3% (2)	3,7% (29)	0,3% (2)									100,0% (784)
D. catholica (6)		5,8% (13)	1,3% (3)			**90,1%** (**201**)										1,3% (3)								1,3% (3)				100,0% (223)
D. cretacea (7)		0,4% (1)	0,7% (2)			2,2% (6)	**88,1%** (**238**)									8,5% (23)												100,0% (270)
D. erucoides subsp. *erucoides* (8)				0,5% (2)	0,9% (4)			**81,1%** (**360**)	4,1% (18)	0,5% (2)	3,6% (16)	0,2% (1)	0,2% (1)	0,5% (2)	0,9% (4)				0,2% (1)		0,2% (1)	0,2% (1)	7,0% (31)					100,0% (444)
D. erucoides subsp. *longisiliqua* (9)				13,7% (66)	0,4% (2)			0,2% (1)	**83,4%** (**403**)		0,2% (1)				1,4% (7)								0,6% (3)					100,0% (484)
D. harra subsp. *harra* (10)	0,9% (3)									**96,7%** (**326**)	0,6% (2)	0,9% (3)										0,9% (3)					100,0% (337)
D. harra subsp. *crassifolia* (11)	1,2% (5)				3,4% (14)					2,9% (12)	**89,9%** (**274**)	1,2% (5)	0,2% (1)									1,2% (5)					100,0% (416)
D. harra subsp. *lagascana* (12)	10,9% (26)				5,9% (14)			0,4% (1)		1,7% (4)	9,2% (22)	**65,7%** (**157**)	2,1% (5)				1,3% (3)	1,3% (3)		1,35 (3)		0,4% (1)						100,0% (239)
D. ibicensis (13)	1,3% (3)				10,7% (25)							1,7% (4)	**70,8%** (**165**)	3,9% (9)	1,3% (3)		2,6% (6)	7,7% (18)										100,0% (233)
D. ilorcitana (14)	0,4% (1)				4,1% (10)						1,2% (3)	1,2% (3)	11,6% (28)	**75,9%** (**183**)	0,4% (1)		1,2% (3)			0,8% (2)	0,4% (1)	2,5% (6)					100,0% (241)
D. muralis subsp. *muralis* (15)	0,6% (1)			1,2% (2)	5,3% (9)			0,6% (1)	5,3% (9)		5,3% (9)		5,3% (9)		**71,0%** (**120**)				0,6% (1)	0,6% (1)		1,8% (3)	2,4% (4)					100,0% (169)
D. muralis subsp. ceratophylla (16)		4,3% (7)	2,5% (4)			3,7% (6)	20,9% (34)									**66,9%** (**109**)										1,2% (2)	0,6% (1)	100,0% (163)
D. ollivieri (17)	1,4% (1)				1,4% (1)					1,4% (1)		1,4% (1)	1,4% (1)	4,3% (3)			**78,3%** (**54**)	7,2% (5)	1,4% (1)	1,4% (1)								100,0% (69)
D. siettiana (18)	3,6% (9)				7,5% (19)							1,6% (4)	3,6% (9)				2,0% (5)	**80,6%** (**204**)		1,2% (3)								100,0% (253)
D. silifolia subsp. *silifolia* (19)					0,8% (2)								0,4% (1)						**89,2%** (**223**)	2,8% (7)	6,4% (22)		0,4% (1)				100,0% (250)
D. silifolia subsp. *bipinnatifida* (20)													0,3% (1)				2,0% (7)	0,3% (1)	7,6% (27)	**83,8%** (**299**)	6,2% (22)						100,0% (357
D. silifolia subsp. *vicentina* (21)																	0,4% (1)		9,3% (26)	9,6% (27)	**80,8%** (**227**)						100,0% (281)
D. simplex (22)	0,8% (3)				1,5% (3)			0,3% (1)		4,3% (17)	0,3% (1)		0,3% (1)	1,3% (5)	0,3% (1)					0,5% (2)		**90,7%** (**360**)						100,0% (397)
D. tenuifolia (23)								16,8% (50)	0,7% (2)					0,7% (2)	1,3% (4)	0,3% (1)			0,3% (1)	0,3% (1)		2,0% (6)	**77,5%** (**231**)					100,0% (298)
D. tenuisiliqua subsp. *tenuisiliqua* (24)		2,8% (14)	2,4% (12)			0,2% (1)										0,2% (1)								**77,8%** (**386**)	9,7% (48)		6,9% (34)	100,0% (496)
D. tenuisiliqua subsp. *rupestris* (25)			7,0% (33)																					9,5% (45)	**64,5%** (**305**)		19,0% (90)	100,0% (473)
D. viminea (26)			0,9% (1)				0,9% (1)									0,9% (1)										**97,3%** (**107**)		100,0% (110)
D. virgata (27)			1,9% (9)													0,9% (4)								7,7% (36)	10,5% (49)		**78,9%** (**367**)	100,0% (465)
Overall																												**80.7%** (**8918**)

Table 6. Cross-validated percentages of correct identification for *D. harra* species classifier at subspecies level. The number of seeds is in parentheses.

Taxa	*D. harra* subsp. *crassifolia*	*D. harra* subsp. *lagascana*	*D. harra* subsp. *harra*	Total
D. harra subsp. *crassifolia*	**96.7% (402)**	1.4% (6)	1.9% (8)	100.0% (416)
D. harra subsp. *lagascana*	9.2% (22)	**88.7% (212)**	2.1% (5)	100.0% (239)
D. harra subsp. *harra*	0.6% (2)	1.2% (4)	**98.2% (331)**	100.0% (337)
Overall				**95.3% (992)**

As expected when subspecies of the same species are compared, the three taxa belonging to *D. harra* were mostly misclassified, while most remaining errors accounted for *D. acris*, *D. brevisiliqua* and *D. simplex*; three southern species with generally smaller seeds, which might have similar dispersal strategies. However, whereas *D. harra* subsp. *harra* and subsp. *crassifolia* did perform satisfactorily (96.7% and 89.9%, respectively), subsp. *lagascana* was less accurate, since only 66% of identified seeds were correct. Only a fraction of its misidentified seeds were mistaken for those of other subspecies of *D. harra* (9.2% as subsp. *crassifolia* and 1.7% as subsp. *harra*). In a separate comparison among the three subspecies of *D. harra*, an overall percentage of correct classification of 95.3% was achieved, but once again, the identification of subsp. *lagascana* (88.7%) was rather lower than the others (Table 6 and Fig. 2).

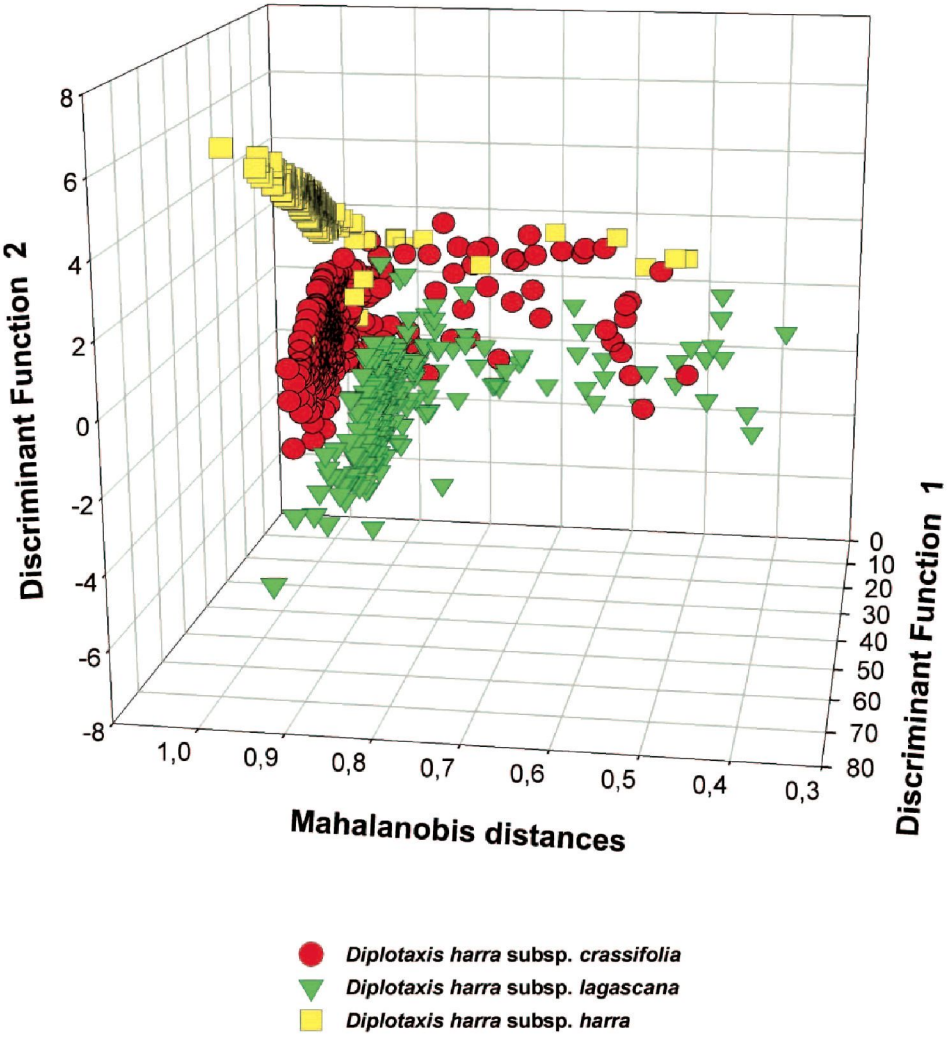

Fig. 2. Graphic representation of the discriminant function scores for *Diplotaxis harra*.

Table 7. Cross-validated percentages of correct identification for *D. muralis* species classifier at subspecies level. The number of seeds is in parentheses.

Taxa	*D. muralis* subsp. *ceratophilla*	*D. muralis* subsp. *muralis*	Total
D. muralis subsp. *ceratophilla*	**100.0% (163)**	0	100.0% (163)
D. muralis subsp. *muralis*	0	**100.0% (169)**	100.0% (169)
Overall			**100.0% (332)**

The North African taxon, *D. muralis* subsp. *ceratophylla* (subgen. *Diplotaxis*), was only correctly identified to the level of 66.9%. A high proportion of errors were due to mistakes for the related species *D. cretacea* (20.9%), and more than 10% were mistaken for species in subgen. *Rhynchocarpum* (*D. assurgens*, 4.3%; *D. berthautii*, 2.5%; *D. catholica*, 3.7%; *D. virgata*, 0.6%), all of them also growing in northern Africa. The seeds of the type subspecies, *D. muralis* subsp. *muralis*, did not perform much better, scarcely achieving 71.0% of correct identification. Only 4.2% of the remaining seeds were mistaken for its close relatives, *D. tenuifolia* (2.4%) and *D. simplex* (1.8%), while most misidentified seeds were wrongly identified as a varied array of taxa in the same subgenus (*D. harra* subsp. *crassifolia*, 5.3%) or in subgen. *Rhynchocarpum* (*D. brevisiliqua*, 5.3%; *D. ibicensis*, 5.3%; *D. erucoides* subsp. *longisiliqua*, 5.3%). Such dispersion in misattribution might be related to the amphidiploid condition of *D. muralis* (n = 21, the only known case of polyploidy in the genus), since polymorphism regularly associated with polyploidy might also be expressed in seed morphology. Surprisingly enough, none of the misidentified seeds of any of the two subspecies of *D. muralis* was mistaken for the other subspecies. Moreover, as shown in Table 7, comparing both taxa separately, the classifier achieved 100% correct identification. Such clear-cut separation is more characteristic of distinct species, rather than of closely related subspecies of the same species and suggests that the taxonomic relationships between them should be reconsidered.

The remaining species in subgen. *Diplotaxis* were more satisfactorily identified. The seeds of *D. simplex* reached 90.7% of the correct identification, with 4.3% of the misidentified seeds attributed to *D. harra* subsp. *harra*, in the same subgenus which is north African. In the case of *D. tenuifolia*, 77.5% of its seeds were well discriminated, with a high proportion of misidentifications being due to confusion with *D. erucoides* subsp. *erucoides* (16.8%), a species belonging to subgen. *Rhynchocarpum*, but with a more similar geographical distribution (mostly central and southern Europe and the Middle East) and probably adapted to more similar habitats. In fact, these are the two species of *Diplotaxis* most frequently considered as weeds (Sans & Masalles, 1994; Eschmann-Grupe *et al.*, 2004). A quite satisfactory correct identification (88.1%) was attained with *D. cretacea* seeds, and the remaining seeds (8.5%) being mostly mistaken for those of *D. muralis* subsp. *ceratophylla*, in the same subgenus. Interestingly, no mistakes occurred between *D. cretacea* (n = 11) and the morphologically similar *D. tenuifolia* (n = 11), which does not support the subordination of the former to the latter proposed by Sobrino Vesperinas (1996). The seeds of *D. viminea*, the only species in the genus that appears to be completely self-fertilising, achieved the highest percentage of correct identification (97.3%). This high identification performance is probably related to the homogeneity regularly associated with autogamy.

In the case of *D. tenuisiliqua* subsp. *rupestris* the low level of successful identification (64.5%) corresponds to the partial confusion with subsp. *tenuisiliqua* (9.5%), but mostly with *D. virgata* (19.0%) and also with *D. berthautii* (7%), both belonging to subgen. *Rhynchocarpum* sect. *Rhynchocarpum* and closely related to *D. tenuisiliqua* subsp. *rupestris*. As for subsp. *tenuisiliqua*, 77.8% of its seeds were correctly determined, with all mistakes occurring with closely related taxa in the same section: *D. tenuisiliqua* subsp. *rupestris* (9.7%), *D. virgata* (6.9%), *D. assurgens* (2.8%) and *D. berthautii* (2.4%). In a separate comparison between the two subspecies of *D. tenuisiliqua*, an identification performance of 92.3% was reached (Table 8 and Fig. 3).

The identification of seeds of all other species in sect. *Rhynchocarpum* was well above 70%. The germplasm of

Table 8. Cross-validated percentages of correct identification for *D. tenisiliqua* species classifier at subspecies level. The number of seeds is in parentheses.

Taxa	*D. tenisiliqua* subsp. *rupestris*	*D. tenisiliqua* subsp. *tenisiliqua*	Total
D. tenisiliqua subsp. *rupestris*	**91.1% (426)**	9.9% (47)	100.0% (473)
D. tenisiliqua subsp. *tenisiliqua*	5.6% (28)	**94.4% (468)**	100.0% (496)
Overall			**92.3% (969)**

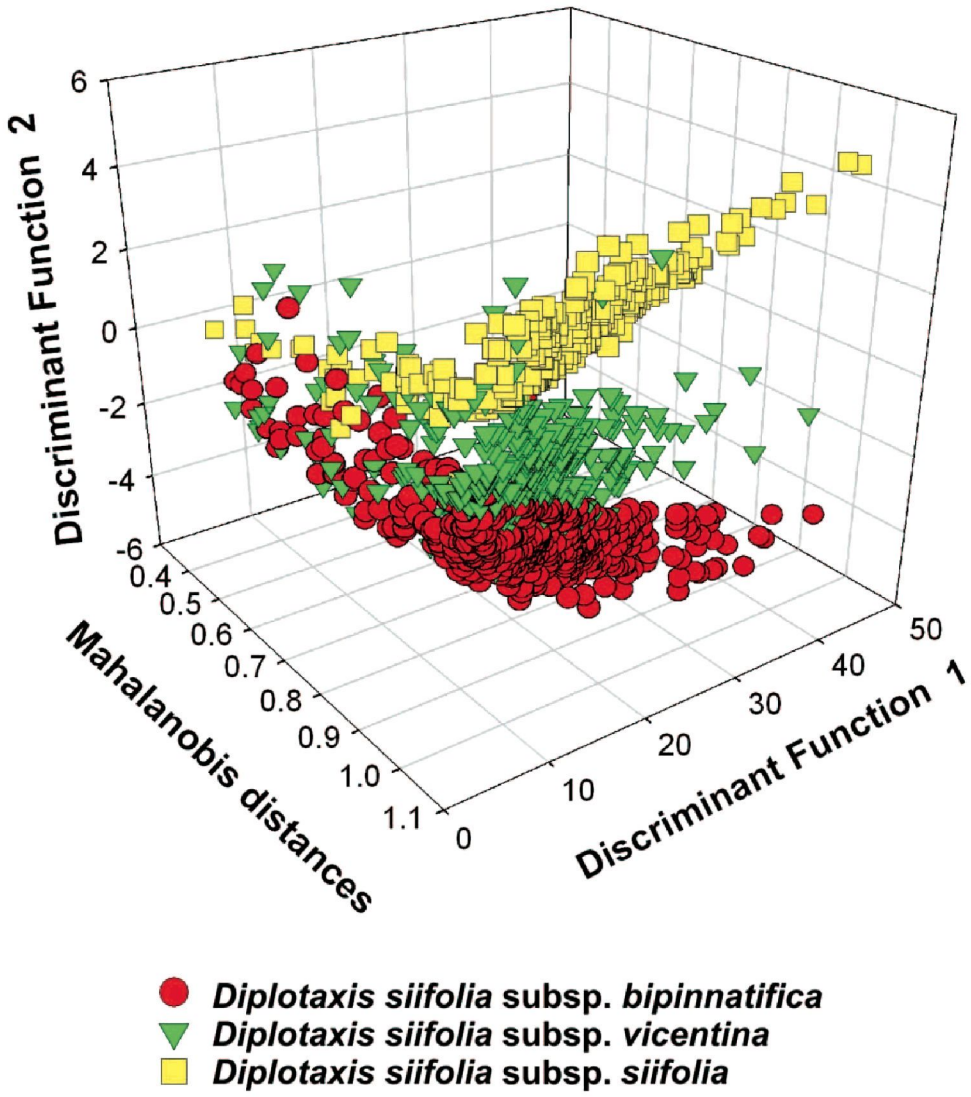

Fig. 3. Graphic representation of the discriminant function scores for *Diplotaxis siifolia*.

D. catholica was quite satisfactorily identified (90.1% of correctly identified seeds), as were those of *D. berthautii* (86.5%), *D. assurgens* (82.4%) and of *D. virgata* (78.9%). Most mistakes occurred among these four species, together with *D. tenuisiliqua*. They are quite related in morphological traits, other than those of the seed, have the same chromosome number (see Table 1) and also share partially similar habitats in the Iberian Peninsula–Morocco region. The germplasm of *D. brachycarpa* was also rather well identified (73.3%). Most errors in this case, however, were due to confusion with *D. erucoides* subsp. *longisiliqua* (22.0%), of section *Heteropetalum*, which in turn reached 83.4% of the correct identification, with 13% mistakes for *D. brachycarpa*. These two taxa belong to different sections of subgen. *Rhynchocarpum*, but grow in a rather limited area in north-eastern Algeria. Since size (Harper *et al.*, 1970) and possibly other seed traits can be affected by selective pressures, convergence might well have caused this similarity between these two taxa growing in similar habitats.

The seeds of *D. siifolia* were also well discriminated, with identification performance between 80.8% and 89.2% for the three subspecies and almost every wrong identification was due to mistakes for one or another of them. Comparing the three taxa separately, the classifier gave an overall correct identification of 92.0% (Table 9), with few mistakes regularly distributed among the subspecies (Fig. 3). These results are in accordance with previous observations that *D. siifolia* is the only *Diplotaxis* species with globose, nearly spherical seeds, more close in shape to those of *Brassica* and other genera in the tribe Brassiceae, a unique feature already pointed out by Bengoechea & Gómez-Campo (1975).

Table 9. Cross-validated percentages of correct identification for *D. siifolia* species classifier at subspecies level. The number of seeds is in parentheses.

Taxa	*D. siifolia* subsp. *bipinnatifica*	*D. siifolia* subsp. *vicentina*	*D. siifolia* subsp. *siifolia*	Total
D. siifolia subsp. *bipinnatifica*	**93.0% (332)**	5.6% (20)	1.4% (5)	100.0% (357)
D. siifolia subsp. *vicentina*	6.4% (18)	**87.5% (246)**	6.0% (17)	100.0% (281)
D. siifolia subsp. *siifolia*	2.4% (6)	2.0% (5)	**95.6% (239)**	100.0% (250)
Overall				**92.0% (888)**

Discrimination of *D. ollivieri* reached 78.3% of correctly identified seeds, with wrongly identified seeds corresponding mostly to all four species of sect. *Heterocarpum* (*D. siettiana*, 7.2%; *D. ilorcitana*, 4.3%; *D. ibicensis*, 1.4%; *D. brevisiliqua*, 1.4%), and the remaining ones to five other taxa. Convergence of seed traits due to similarity of habitats, again, might be the cause for the misidentification among these taxa. However, *D. ollivieri* has never been investigated for chromosome number or molecular markers, and its morphological affinities are poorly understood because descriptions are based on scarce availability of herbarium specimens. Its taxonomic position in sect. *Rhynchocarpum* is therefore weakly sustained and somewhat uncertain, and does not seem to be supported by seed morpho-colorimetric characters.

Seeds of all four species of section *Heterocarpum* achieved percentages of correct identification ranging from 70.8% (*D. ibicensis*) to 83.2% (*D. brevisiliqua*). Furthermore, wrongly identified seeds have generally been mistaken for seeds of other species of the same section.

The only two taxa in section *Heteropetalum* are the two subspecies of *D. erucoides* and both turned out to be well identified in more than 80% of cases. Most errors were due to misclassification for species from more similar habitats. Seeds of *D. erucoides* subsp. *erucoides* were mainly misidentified for *D. tenuifolia* (7.0%) and *D. harra* subsp. *crassifolia* (3.6%), both of them with brochidodromously veined petals and a native habitat in central and southern Europe, as *D. erucoides* subsp. *erucoides*; and only 4.1% of the seeds was mistaken for those of subsp. *longisiliqua*, which grows in northern Algeria. On the other hand, some of the seeds of *D. erucoides* subsp. *longisiliqua* were incorrectly identified as *D. brachycarpa* (13.7%). Misidentifications between the two subspecies; however, occurred only in a very small proportion; both were in fact perfectly discriminated when compared separately (Table 10).

Infrageneric classification

The distribution of *Diplotaxis* taxa in the three-dimensional space, was determined by the first three discriminant functions (DF) derived from morpho-colorimetric parameters obtained from seed image analysis (Fig. 4). Two major groups of taxa are recognisable graphically.

This first cloud includes most taxa of sect. *Rhynchocarpum*, with the exceptions of *D. siifolia* subspecies, *D. ollivieri* (both of questionable taxonomic position) and *D. brachycarpa*, as well as a subset constituted by two additional taxa belonging to subgen. *Diplotaxis*: *D. cretacea* and *D. muralis* subsp. *ceratophylla*. The latter appears quite distant from subsp. *muralis*, which in turn is situated in the other set of taxa, among the remaining subgen. *Diplotaxis*. The fact that the two subspecies of *D. muralis* appear so far apart reinforces the idea that, at least on the basis of morpho-colorimetric traits of seeds, these two taxa might well be considered as separate species. The striking placement of *D. cretacea* cannot be explained on the basis of any affinity other than seed morphology, since according to plant morphology, isozymes, chromosome numbers and molecular markers, it is clearly much closer to other taxa of the subgen. *Diplotaxis*.

The second major group is located basically in the opposite corner of the graph. It consists of all other taxa belonging to subgen. *Diplotaxis*, together with those of sect. *Heteropetalum*, sect. *Heterocarpum* and subgen. *Hesperidium*, plus the previously mentioned *D. brachycarpa*, *D. siifolia* subspecies and *D. ollivieri* of sect. *Rhynchocarpum*. The taxa belonging to the type subgenus were all found to be included in a clade within the Rapa–Oleracea lineage

Table 10. Cross-validated percentages of correct identification for *D. erucoides* species classifier at subspecies level. The number of seeds is in parentheses.

Taxa	*D. erucoides* subsp. *longisiliqua*	*D. erucoides* subsp. *erucoides*	Total
D. erucoides subsp. *longisiliqua*	**100.0% (483)**	0	100.0% (483)
D. erucoides subsp. *erucoides*	0.02% (1)	**99.8% (443)**	100.0% (444)
Overall			**99.9% (927)**

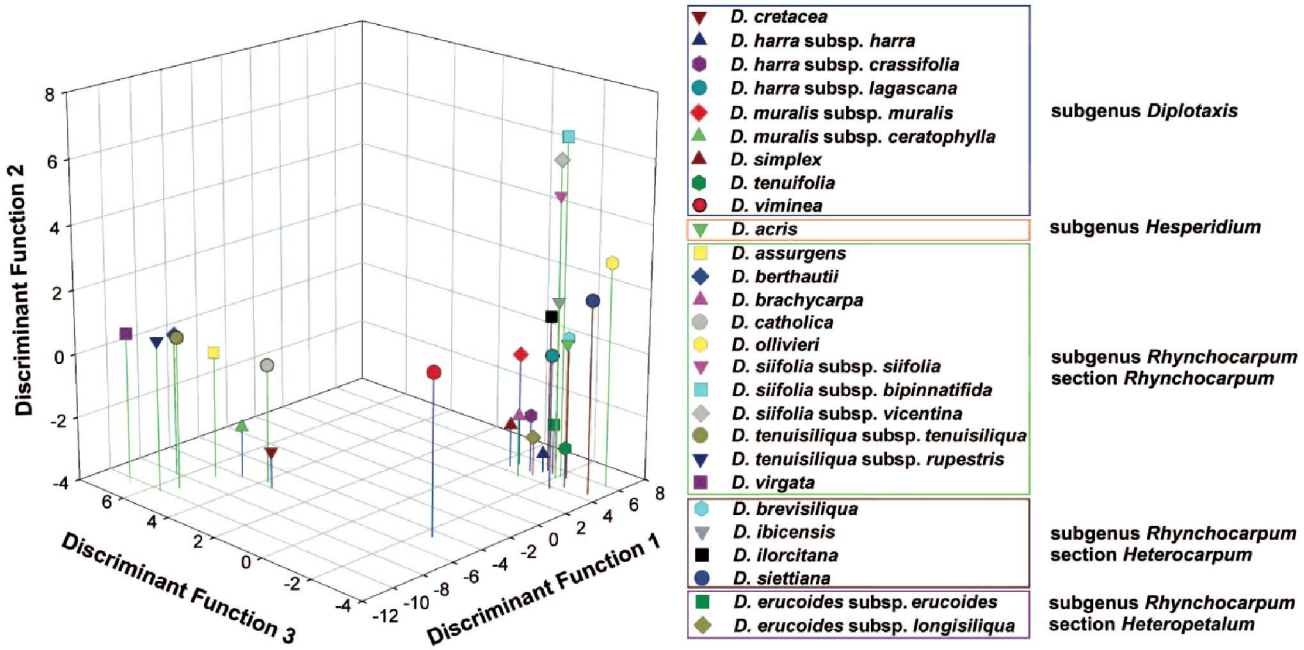

Fig. 4. Graphic representation of the discriminant function scores for the studied *Diplotaxis taxa*.

(Warwick *et al.*, 1992) and form a rather compact set, with negative values of DF2, except for *D. viminea*, which due to the self-fertilisation syndrome of *D. viminea*, makes it morphologically different from its closest relatives, in many aspects, and seed morpho-colorimetric characteristics may well be among them.

Intermixed among taxa of subgen. *Diplotaxis* *D. acris* also appears, whose separate status as subgen. *Hesperidium* does not seem to be supported, therefore, by seed characters.

The two subspecies of *D. erucoides* were plotted close to each other, but also among taxa of subgen. *Diplotaxis*, suggesting that sect. *Heteropetalum*, in terms of seed morpho-colorimetric characteristics, is closer to the type subgen. than to subgen. *Rhynchocarpum*. Similarly, both taxa in this section, as well as those of subgen. *Diplotaxis*, were found by Warwick *et al.* (1992) to belong to the same clade, the Rapa–Oleracea lineage. It should be noted that subgen. *Diplotaxis*, subgen. *Hesperidium* and sect. *Heteropetalum* also share the brochidodromous venation of petals, in contrast with the cladodromous to eucamptodromous petals of the remaining sections of subgen. *Rhynchocarpum*.

Sect. *Heterocarpum* is well supported on molecular grounds (Warwick *et al.*, 1992) and appears to be confirmed as a taxonomic unit also on the basis of seed morphology, since its four species (*D. brevisiliqua*, *D. ibicensis*, *D. ilorcitana* and *D. siettiana*) form a quite concise group positioned rather close to subgen. *Diplotaxis*. Subgenus *Rhynchocarpum* is also represented in this second major cloud by sect. *Heteropetalum* (discussed above), *D. ollivieri* (of uncertain taxonomic position), *D. brachycarpa* and *D. siifolia*. The three subspecies of *D. siifolia* (sect. *Rhynchocarpum*) can still be found somewhat above sect. *Heterocarpum*. The odd position of this species might well be related to the spherical shape of its seeds, unique in the genus. Besides, *D. siifolia* has n = 10 chromosomes (a number only shared with the not closely related *D. viminea*) instead of n = 9 chromosomes, as the rest of sect. *Rhynchocarpum*. Several authors (Gómez-Campo & Tortosa, 1974; Takahata & Hinata, 1983, 1986) have pointed out the difficulty in accommodating *D. siifolia*, on morphological grounds, in *Diplotaxis* as well as its affinities with *Brassica* or *Erucastrum*. The placement of *D. brachycarpa* close to subgen. *Diplotaxis* seems to be solely founded on seed morphology, since according to other morphological traits, chromosome numbers and molecular markers, it is clearly much closer to subgen. *Rhynchocarpum*, though it was found to form a separate subclade within the Nigra lineage by Warwick *et al.* (1992).

Conclusions

This work represents the first approach to investigate taxonomic relationships within the genus *Diplotaxis* by seed phenetic characterisation using a seed image analysis system.

The results obtained by the implementation of the general database for this genus and the elaboration of dedicated seed classifiers for the subspecies groups within this genus, prove once again how image analysis techniques can be considered a useful tool in taxonomic studies. In this case, the application of this innovative kind of identification system allowed us to discriminate among most species and

subspecies with a considerably high percentage of correct identification, as well as supporting the most recently proposed infrageneric grouping. In particular, the consistency of section *Heterocarpum*, the particular position of section *Heteropetalum*, and the rather isolated situation of *D. siifolia* with respect to the rest of the genus are supported by this analysis.

A certain amount of misidentified seeds was also found, as should always be expected considering that samples collected in the wild are heterogeneous, and consequently may have different levels of maturation, may show intrapopulation genetic variation, and probably are subject to any other endogenous and exogenous interactions or sources of natural heterogeneity, as well as to plausible phenomena of convergence among species adapted to similar environmental conditions.

Acknowledgements

The authors thank the "Regione Autonoma Sardegna" for the support and the realization of this work, on the basis of the Legge Regionale 7 agosto 2007, n. 7 "Promozione della ricerca scientifica e dell'innovazione tecnologica in Sardegna", pubblica selezione per il conferimento di borse di ricerca destinate a giovani ricercatori – Bando 2008.

References

AMIN, A. 1972. IOPB Chromosome number reports XXXVIII. *Taxon* **21**, 679–684.

BACCHETTA, G., FENU, G., GRILLO, O., MATTANA, E. & VENORA, G. 2011*a*. Taxonomic identification by image analysis technique of seeds in the *Astragalus* Sect. *Melanocercis* Bunge (Fabaceae) species of Sardinia. *Annales Botanici Fennici* **48**, 449–454.

BACCHETTA, G., GARCÌA, P.E., GRILLO, O., MASCIA, F. & VENORA, G. 2011*b*. Seed image analysis provides evidence of taxonomical differentiation within the *Lavatera triloba* aggregate (Malvaceae). *Flora* **206**, 468–472.

BACCHETTA, G., GRILLO, O., MATTANA, E. & VENORA, G. 2008. Morpho-colorimetric characterization by image analysis to identify diaspores of wild plant species. *Flora* **203**, 669–682.

BENGOECHEA, G. & GÓMEZ-CAMPO, C. 1975. Algunos caracteres de la semilla en la tribu Brassiceae. *Anales del Instituto Botanico A.J. Cavanilles* **32**, 793–841.

DANA, W. & IVO, W. 2008. Computer image analysis of seed shape and seed color for flax cultivar description. *Computer and Electronics in Agriculture* **61**, 126–135.

ESCHMANN-GRUPE, G., HURKA, H. & NEUFFER, B. 2003. Species relationships within *Diplotaxis* (Brassicaceae) and the phylogenetic origin of *D. muralis*. *Plant Systematics and Evolution* **243**, 13–29.

ESCHMANN-GRUPE, G., NEUFFER, B. & HURKA, H. 2004. Extent and structure of genetic variation in two colonising *Diplotaxis* species (Brassicaceae) with contrasting breeding systems. *Plant Systematics and Evolution* **244**, 31–43.

FAGUNDEZ, J. & IZCO, J. 2004. Seed morphology of *Calluna* Salisb. (Ericaceae). *Acta Botanica Malacitana* **29**, 215–220.

FERNANDES, A. & QUEIRÓS, M. 1970–71. Sur la caryologie de quelques plantes recoltées pendant la IIIème réunion de botanique peninsulaire. *Memória de Sociedade Broteriana* **21**, 343–385.

FISHER, R.A. 1936. The use of multiple measurements in taxonomic problems. *Annals of Eugenics* **7**(2), 179–188.

FISHER, R.A. 1940. The precision of discriminant functions. *Annals of Eugenics* **10**(4), 422–429.

GÓMEZ-CAMPO, C. 1980. Studies on Cruciferae, V. Chromosome numbers for twenty-five taxa. *Anales del Instituto Botánico A.J. Cavanilles* **35**, 177–182.

GÓMEZ-CAMPO, C. & TORTOSA, M.E. 1974. The taxonomic and evolutionary significance of some juvenile characters in the tribe Brassiceae. *Botanical Journal of the Linnean Society* **69**, 105–124.

GÓMEZ-CAMPO, C. & MARTÍNEZ-LABORDE, J.B. 1998. Reajustes taxonómicos y nomenclaturales en la tribu Brassiceae (Cruciferae). *Anales del Jardín Botánico de Madrid* **56**, 379–381.

GRANITTO, P.M., GARRALDA, P.A., VERDES, P.F. & CECCATO, H.A. 2003. Boosting classifiers for weed seeds identification. *Journal of Computer Science and Technology* **3**, 34–39.

GRILLO, O., MATTANA, E., VENORA, G. & BACCHETTA, G. 2010. Statistical seed classifiers of 10 plant families representative of the Mediterranean vascular flora. *Seed Science and Technology* **38**, 455–476.

GRILLO, O., MICELI, C. & VENORA, G. 2011. Computerised image analysis applied to inspection of vetch seeds for varietal identification. *Seed Science and Technology* **39**, 490–500.

HARBERD, D.J. 1972. A contribution to the cyto-taxonomy of *Brassica* (Cruciferae) and its allies. *Botanical Journal of the Linnean Society* **65**, 1–23.

HARBERD, D.J. 1976. Cytotaxonomic studies of *Brassica* and related genera. In: VAUGHAN, J.G., MACLEOD, A.J. & JONES, B.M.G., Eds., *The Biology and Chemistry of the Cruciferae*. Academic Press, London, pp. 47–68.

HARBERD, D.J. & MCARTHUR, E.D. 1972. The chromosome constitution of *Diplotaxis muralis* (L.) DC. *Watsonia* **9**, 131–135.

HARPER, J.L., LOVELL, P.H. & MOORE, K.G. 1970. The shapes and sizes of seeds. *Annual Review of Ecology and Systematics* **1**, 327–356.

KILIÇ, K., BOYACI, I.H., KOKSEL, H. & KUSMENOGLU, I. 2007. A classification system or beans using computer vision system and artificial neural networks. *Journal of Food Engineering* **78**, 897–904.

LIAO, K., PAULSEN, M.R. & REID, J.F. 1994. Real-time detection of colorimetric and surface defects of maize kernels using machine vision. *Journal of Agricultural Engineering Research* **59**, 263–271.

MARTIN, A.C. 1946. The comparative internal morphology of seeds. *American Midland Naturalist Journal* **36**, 513–660.

MARTÍN, J.P. & SÁNCHEZ-YÉLAMO, M.D. 2000. Genetic relationships among species of the genus *Diplotaxis* (Brassicaceae) using inter-simple sequence repeat markers. *Theoretical and Applied Genetics* **101**, 1234–1241.

MARTÍNEZ-LABORDE, J.B. 1988. Estudio sistemático del género *Diplotaxis* DC. (Cruciferae-Brassiceae). PhD dissertation, Universidad Politécnica de Madrid, Madrid, ES.

MARTÍNEZ-LABORDE, J.B. 1991. Two additional species of *Diplotaxis* (Cruciferae, Brassiceae) with n = 8 chromosomes. *Willdenowia* **21**, 63–68.

PIGNONE, D. 1997. Present status of rocket genetic resources and conservation activities. In: PADULOSI, S. & PIGNONE, D., Eds., *Rocket: A Mediterranean Crop for the World*. IPGRI, Rome, pp. 2–12.

PIMPINI, F. & ENZO, M. 1997. Present status and prospects for rocket cultivation in the Veneto region. In: PADULOSI, S. & PIGNONE, D., Eds., *Rocket: a Mediterranean Crop for the World*. IPGRI, Rome, pp. 51–66.

PRADHAN, A.K., PRAKASH, S., MUKHOPADHYAY, A. & PENTAL, D. 1992. Phylogeny of *Brassica* and allied genera based on variation in chloroplast and mitochondrial DNA patterns: molecular and taxonomic classifications are incongruous. *Theoretical and Applied Genetics* **85**, 331–340.

PRAKASH, S., TAKAHATA, Y., KIRTI, P. & CHOPRA, V. 1999. Cytogenetics. In: GÓMEZ-CAMPO C., Ed., *Biology of Brassica Coenospecies*. Elsevier, Amsterdam, pp. 59–106.

RUSTAN, Ø.H. 1996. Revision of the genus *Diplotaxis* (Brassicaceae) in the Cape Verde Islands, W Africa. *Nordic Journal of Botany* **16**, 19–50.

SÁNCHEZ-YÉLAMO, M.D. & MARTÍNEZ-LABORDE, J.B. 1991. Chemotaxonomic approach to *Diplotaxis muralis* (Cruciferae: Brassiceae) and related species. *Biochemical Systematics and Ecology* **19**, 477–482.

SANS, F.X. & MASALLES, R.M. 1994. Life-history variation in the annual arable weed *Diplotaxis erucoides* (Cruciferae). *Canadian Journal of Botany* **72**, 10–19.

SHAHIN, M.A. & SYMONS, S.J. 2003*a*. Lentil type identification using machine vision. *Canadian Biosystems Engineering* **45**, 3.5–3.10.

SHAHIN, M.A. & SYMONS, S.J. 2003*b*. Colour calibration of scanners for scanner independent grain grading. *Cereal Chemistry* **80**, 285–289.

SMYKALOVA, I., GRILLO, O., BJELKOVA, M., HYBL, M. & VENORA, G. 2011. Morpho-colorimetric traits of *Pisum* seeds measured by an image analysis system. *Seed Science and Technology* **39**, 612–626.

SOBRINO VESPERINAS, E. 1996. Posición taxonómica de *Diplotaxis cretacea* Kotov (Cruciferae). *Anales del Jardín Botánico de Madrid* **54**, 182–188.

SPSS, 1999. *Base 10.0 Application Guide*. Prentice Hall, New Jersey.

STEARN, W.T. 1980. *Botanical Latin*. David & Charles Publ., London.

TAKAHATA, Y. & HINATA, K. 1978. A description of the genetic stocks in subtribe Brassicinae by chromosome numbers and numerical characters. *Cruciferae Newsletter* **3**, 47–51.

TAKAHATA, Y. & HINATA, K. 1983. Studies on cytodemes in subtribe Brassicinae (Cruciferae). *Tohoku Journal of Agricultural Research* **33**, 111–124.

TAKAHATA, Y. & HINATA, K. 1986. Consideration of the species relationships in subtribe Brassicinae (Cruciferae) in view of cluster analysis of morphological characters. *Plant Species Biology* **1**, 79–88.

VENORA, G., GRILLO, O., SHAHIN, M.A. & SYMONS, S.J. 2007. Identification of Sicilian landraces and Canadian cultivars of lentil using image analysis system. *Food Research International* **40**, 161–166.

VENORA, G., GRILLO, O. & SACCONE, R. 2009*a*. Quality assessment of durum wheat storage centres in Sicily: Evaluation of vitreous, starchy and shrunken kernels using an image analysis system. *Journal of Cereal Science* **49**, 429–440.

VENORA, G., GRILLO, O., RAVALLI, C. & CREMONINI, R. 2009*b*. Identification of Italian landraces of bean (*Phaseolus vulgaris* L.) using an image analysis system. *Scientia Horticulturae* **121**, 410–418.

WARWICK, S.L. & ANDERSON, J.K. 1997. Isozyme analysis of parentage in allopolyploid *Diplotaxis muralis* (L.) DC. (Brassicaceae). *Cruciferae Newsletter* **19**, 35–36.

WARWICK, S.L. & BLACK, L.D. 1991. Molecular systematics of *Brassica* and allied genera (Subtribe Brassicinae, Brassiceae) chloroplast genome and cytodeme congruence. *Theoretical and Applied Genetics* **82**, 81–92.

WARWICK, S.L., BLACK, L.D. & AGUINAGALDE, I. 1992. Molecular systematics of *Brassica* and allied genera (Subtribe Brassicinae, Brassiceae) chloroplast DNA variation in the genus *Diplotaxis*. *Theoretical and Applied Genetics* **83**, 839–850.

WARWICK, S.L., FRANCIS, A. & AL-SHEHBAZ, I.A. 2006. Brassicaceae: species checklist and database on CD-rom. *Plant Systematics and Evolution* **259**, 249–258.

WERKER, E. 1997. *Seed Anatomy. Encyclopedia of Plant Anatomy,* Vol. 10. Borntraeger, Berlin.